

Spam Filtering

This document contains the following sections:

- [Mail Filter](#)
- [The "mailto:" Tag](#)
- [Spam Resistant Email List](#)
- [Spam Filter through the Portal](#)
(<http://www.cmu.edu/computing/doc/email/spam/portal/index.html>)
- [How the Andrew Spam Filter Works](#)

For information related to this topic refer to:

- [Spam Filtering](#) (<http://www.cmu.edu/computing/email/spam-filtering.html>)
- [Using Sieve for Email Filtering](#)
(<http://www.cmu.edu/computing/doc/email/sieve/index.html>)

Mail Filter

A Mail Filter, released in spring 2003, gives Cyrus e-mail users the ability to significantly reduce the amount of spam or junk mail messages in their INBOX folder through the following three functions:

- Enable Spam filtering. This feature allows you to turn spam filtering on and off. Spam will be filed into a folder, which you specify in a dialogue box (e.g., INBOX.spam). Instructions on how to create folders will be available on the page.
- Create and Modify Accept List. The Accept List is a list of addresses or domains from which you always want to receive e-mail messages, even if the spam filter thinks that they could be spam. You can add or remove addresses (e.g., ju33@andrew.cmu.edu) or domains (e.g., hotmail.com) using the tools on the page.
- Create and Modify Filter List. The Filter List is a list of addresses and domains which you always want to treat as spam. Thus, if you always want to filter e-mail from ju33@andrew.cmu.edu or from an AOL user (aol.com) into your spam folder, you can add those to your Filter List and they'll always be filed into your spam folder.

To access and enable the Mail Filter, visit the [My Andrew Web Service \(https://www.cmu.edu/myandrew/\)](https://www.cmu.edu/myandrew/) and click on the link for "Mail Filtering and Vacation".

Last Updated: 7/13/04

The "mailto:" Tag

Spam Robots or Spiders "crawl" the Internet searching for "mailto:" tags. To create a clickable "mailto:" hyperlink that is spam resistant, use the following JavaScript on your web pages.

```
<SCRIPT LANGUAGE="JavaScript">
<!-- Begin
user = "myusername";
site = "myhost.com";
document.write('<a href="mailto:' + user + '@' + site + ">');
document.write('Email me!' + '</a>');
// End -->
</SCRIPT>
```

This will result in a "mailto:" like this: [Email me!](mailto:post+org.acs.documentation@andrew.cmu.edu)
(<mailto:post+org.acs.documentation@andrew.cmu.edu>)

Steps to create a clickable "mailto:" hyperlink:

1. Copy and insert the above JavaScript into your html source code at the mailto: insertion point.
Note: Some html authoring tools may require you to use an automated tool for inserting the JavaScript (e.g., Insert > Script Object in Dreamweaver).
2. Edit the script to include the user name and mail host.
For example, suppose email should be sent to jdoe@andrew.cmu.edu. Within the script, the lines that read:
user = "myusername";
site = "myhost.com";
must be changed to:
user = "jdoe";
site = "andrew.cmu.edu";
3. Save your html file. The JavaScript will generate a clickable link that reads "Email me!".

Last Updated: 7/13/04

Spam Resistant Email List

If you must maintain a web page of email addresses, restrict it to one page. You can limit the number of SpamBots from indexing the page with the following meta tag:

"meta name="Robots" content="noindex,nofollow"

Most SearchBots, SpamBots & web crawlers ignore pages with this tag.

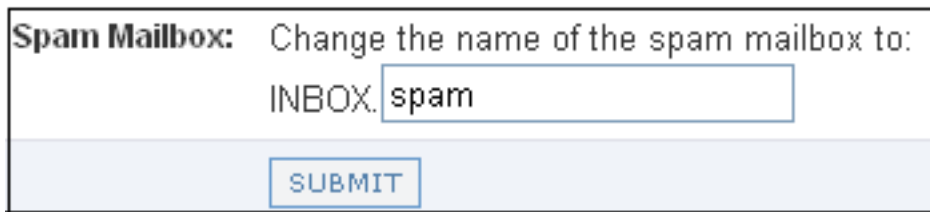
Last Updated: 7/13/04

Last Reviewed 5/21/09

The INBOX.spam Mailbox

When the Andrew Spam Filter is enabled, a mailbox named INBOX.spam is automatically created for you. If you chose the option to store spam, mail that has been flagged as spam will be filed into this mailbox. If you wish to store your spam in another mailbox, follow these steps:

1. If you haven't already created a mailbox in which you want to store your spam, launch the email client that you use to read your mail (e.g., Outlook, Entourage, etc.) and create a mailbox within your Cyrus INBOX; name the folder appropriately (e.g., junkmail, etc.).
2. From the [Carnegie Mellon portal \(http://my.cmu.edu/\)](http://my.cmu.edu/), select the **COMPUTING** tab.
3. Select **Manage Email** and then **spam**.
4. In the Spam Mailbox area, enter the name of the mailbox that you created in step 1 above.



Spam Mailbox: Change the name of the spam mailbox to:
INBOX.

5. Click **Submit** to save your changes.

Last Updated: 5/21/09

Enable Spam Filter Using Portal

Follow these steps to enable the Andrew Spam Filter:

1. From the [Carnegie Mellon portal \(https://my.cmu.edu/\)](https://my.cmu.edu/), select the **COMPUTING** tab.
2. Select **Manage Email** and then select **spam**.
3. In the "**Turn on/off Spam Filtering**" area select from the following options:
 - Turn OFF Spam Filtering
Note: If you select this option you WILL RECEIVE ALL email, this includes mail that our servers score as spam.
 - Turn ON Spam Filtering and discard Spam and mail from addresses on the Filter Mail list below
Note: If you have listed any addresses in the spam filter's "Accept Mail" list, **DO NOT use this option**. This option DOES NOT consider that list before discarding spam.
 - Turn ON Spam Filtering, apply the Accept & Filter Mail lists and STORE spam in the spam mailbox designated below
Note: If you select this option, email that is scored by our mail servers as spam AND mail that YOU designate as spam will be stored in the designated mailbox. Email from addresses that are included in the "Accept Mail" list WILL be delivered to you.

Spam

Andrew (Cyrus) Email Spam Filtering

Turn OFF Spam Filtering
 Turn ON Spam Filtering and discard Spam and mail from addresses on Mail list below
 Turn ON Spam Filtering, apply the Accept & Filter Mail lists and STORE the spam mailbox designated below.

Accept Mail: Accept incoming mail from these addresses:

Accept Mail List:

Filter Mail: Filter incoming mail from these addresses:

Filter Mail List:

RobDavis@theentertainmentinsider.net
 webmaster@iparenting.com
 JenniferJones@theentertainmentinsider.net

Spam Mailbox: Change the name of the spam mailbox to:

INBOX.

4. Click **Submit**.

IMPORTANT! The Andrew Mail Filter analyzes mail headers and body text to identify mail that is most likely spam. However, due to the varying content and origin of mail messages, legitimate mail may be flagged as spam.

(See [How the Andrew Spam Filter Works](#) for more information on the methods for flagging mail as spam.)

Last Updated: 5/21/09

About the Accept and Filter Mail Lists

The Accept list and Filter list options are tools for customizing your Spam Filter.

- The **Accept Mail List** allows you to identify email addresses or domains that you always want to accept mail from even if the Andrew Spam Filter would normally filter the mail. The mail that is received from the addresses or domains on the Accept Mail List will be delivered to your INBOX.
IMPORTANT! The Accept Mail list is NOT REFERENCED when the option to DISCARD SPAM is selected.
- The **Filter Mail List** allows you to specify email addresses and domains that you NEVER want to accept mail from. Mail received from email addresses or domains on the Filter list will automatically be stored or discarded depending on the filter option that you selected.

To use the Accept list and Filter list features, you must first build a list of email addresses or domains and then enable the Accept list or Filter list. The following sections include instructions for these tasks.

Last Updated: 5/21/09

What is a Domain?

Domains are sites from which e-mail messages originate. In an e-mail address, everything to the right of the "@" symbol is the domain. Some examples of domains include:

- hotmail.com
- andrew.cmu.edu
- pitt.edu
- aol.com

For example: If you have a domain in your Filter list of "cmu.edu," any e-mail sent from User@cmu.edu would be filtered, but email sent from "User@andrew.cmu.edu" or from "User@spamsender.cmu.edu" would not. The domain must match exactly what is on the right side of the @ sign. This is done to prevent accidental filtering of email that partially matches the domain address.

Last Updated: 5/21/09

Configuring the Accept Mail List

Unless you have selected the option to DISCARD SPAM, email addresses and domains that you add to the Accept list will always be delivered to your INBOX. This is true even if they are flagged as spam by the Andrew Spam Filter.

IMPORTANT! The Accept Mail list is NOT REFERENCED when the option to DISCARD SPAM is selected.

Some addresses/domains you may want to consider adding to your Accept list include:

- Email addresses that are an exception to a domain that you added to your Filter list
If you get a lot of spam from Hotmail accounts, you can add the hotmail.com domain to your Filter list. However, if you want to receive mail from a friend's Hotmail account, add their e-mail address to your Accept list as an exception (e.g., jdoe@hotmail.com).
- Mailing list addresses
Messages from mailing lists are often filtered by the Andrew Spam Filter because the TO: address does not match your email address and because the "subscribe/unsubscribe" instructions included in the body are also typically found in unsolicited mail messages.
- Email addresses for mail that is consistently filed in your spam folder.
Because of the special characters used in international mail, some international students have found that legitimate email is being filed in their spam folder. By identifying the email address in the Accept list, this mail will automatically be delivered to your Inbox.

To add an email address or domain to the Accept Mail List:

1. From the [Carnegie Mellon portal \(http://my.cmu.edu/\)](http://my.cmu.edu/), select the **COMPUTING** tab.
2. Select **Manage Email** and then **spam**.
3. Type the email address or domain name (e.g., jdoe@hotmail.com, hotmail.com) in the **Accept Mail** field.

Spam

Andrew (Cyrus) Email Spam Filtering

Turn OFF Spam Filtering
 Turn ON Spam Filtering and discard Spam and mail from addresses on the Filter Mail list below
 Turn ON Spam Filtering, apply the Accept & Filter Mail lists and STORE spam in the spam mailbox designated below.

Accept Mail: Accept incoming mail from these addresses:

Accept Mail List: juser@connecttime.com

4. Click **Add**. The address or domain that you entered is added to the Accept Mail List area.
5. Click **Submit** to save your changes.

To remove an email address or domain from the Accept list:

1. From the [Carnegie Mellon portal \(http://my.cmu.edu/\)](http://my.cmu.edu/), select the **Computing** tab.
2. Select **Manage Email** and then **spam**.
3. In the **Accept Mail List** area, select the e-mail address or domain you want to remove.
4. Click **Remove**. The address or domain that you selected is removed from the Accept List.
5. Click **Submit** to save your changes.

Last Updated: 6/3/09

Configuring the Filter Mail List

The email addresses and domains that you add to the Filter Mail List will be filtered into your spam folder unless the address is also on your Accept list. (The Accept list always takes precedence over the Filter list.) Messages received from the identified email addresses or domains are automatically filed as spam, even if they are not flagged as spam by the Andrew Spam Filter.

To add an e-mail address or domain to the Filter list:

1. From the [Carnegie Mellon portal \(http://my.cmu.edu/\)](http://my.cmu.edu/), select the **COMPUTING** tab.
2. Select **Manage Email** and then **spam**.
3. Type the e-mail address or domain name (e.g., jdoe@hotmail.com, hotmail.com) in the **Filter Mail** field.

Filter Mail: Filter incoming mail from these addresses:

Filter Mail List:

- RobDavis@theentertainmentinsider.net
- webmaster@iparenting.com
- JenniferJones@theentertainmentinsider.net

Spam Mailbox: Change the name of the spam mailbox to:

4. Click **Add**. The address or domain that you entered is added to the Filter Mail List area.
5. Click **Submit** at the bottom of the Advanced Features area to save your changes.

To remove an email address or domain from the Filter list:

1. From the [Carnegie Mellon portal \(http://my.cmu.edu/\)](http://my.cmu.edu/), select the **COMPUTING** tab.
2. Select **Manage Email** and then **spam**.
3. In the **Filter Mail List** area, select the e-mail address or domain you want to remove.
4. Click **Remove**. The address or domain that you selected is removed from the Filter List.
5. Click **Submit** to save your changes.

Last Updated: 5/21/09

How the Andrew Spam Filter Works

The Andrew Spam Filter is implemented using a product called [PureMessage by Sophos](http://www.sophos.com/products/pm/) (<http://www.sophos.com/products/pm/>). PureMessage is a mail filter that uses a rule base to perform tests on e-mail and identify spam. The PureMessage software works in conjunction with our Cyrus mail server and performs these tests on all mail received by the server. However, mail sent through the Andrew outgoing mail server (SMTP.ANDREW.CMU.EDU) is NOT scanned by PureMessage.

Note: Mail sent via the Andrew outgoing mail server that must go out to another server before being delivered to the Cyrus server will be scanned by the Andrew Spam Filter. For example, if you send mail using the Andrew outgoing mail server to a "user@cmu.edu" address, that mail is sent to the cmu.edu server and then to the Cyrus mail server. Because the Cyrus mail server receives the mail from a server other than the Andrew outgoing mail server, that mail will be evaluated and scored by the Andrew Spam Filter.

Following is an overview of how the Andrew Spam Filter handles your incoming mail.

1. All mail received by the Cyrus mail server is scanned by the Andrew Spam Filter.

All incoming e-mail received by the Cyrus mail server is scanned for signs that it may be spam. This scan is performed regardless of whether you have enabled the spam filter. The Andrew Spam Filter contains a series of conditions or rules that it checks against each piece of mail. Remember, mail sent through the Andrew outgoing mail server (SMTP.ANDREW.CMU.EDU) is NOT scanned or "scored".

2. The Andrew Spam Filter offers a percentage score for your email.

Each condition or rule within the Andrew Spam Filter has a grade. The total score for a piece of e-mail is the sum of the grades for each of the conditions that the mail matches. PureMessage then takes this score and converts it into a percentage likelihood that the message is spam.

3. If an e-mail has a percentage of 50 or higher, a line containing the score and reason for the overall score is inserted in the header of the "spam" message.

Following is an example of the information that might be added to the header of your mail message. The information within parenthesis includes the reasons for the overall score for the message (in this case, 92%).

```
X-Spam-Warning: 92% (URI_CLASS_HEALTH_DOMAIN 8, CTYPE_JUST_HTML 0.848, LIMITED_TIME_ONLY 0.461, BIG_FONT 0.146, CLICK_HERE_LINK 0.131, HTML_50_70 0.092, CLICK_BELOW 0.089, __CTYPE_IS_HTML 0, __UNUSABLE_MSGID 0, __CLICK_BELOW 0, __CLICK_HERE_LINK 0, __TAG_EXISTS_BODY 0, __MIME_HTML 0, HTML_FONT_COLOR_YELLOW 0, __MIME_HTML_ONLY 0, __TAG_EXISTS_HTML 0, __TO_MALFORMED_2 0, __MIME_VERSION 0, __EVITE_CTYPE 0, __CT 0, __CTYPE_HTML 0)
```

4. Depending on which spam filter option you selected, mail messages with a percentage less than 50 are delivered to your INBOX. Depending on which spam option you chose, the Andrew Spam Filter either discards or files e-mail messages with an overall percentage score of 50 or greater into your INBOX.spam folder (or another folder that you designated) with the FOLLOWING EXCEPTIONS:

- E-mail sent using the Andrew outgoing mail server (SMTP.ANDREW.CMU.EDU) is NOT scanned by the Andrew Spam Filter. Therefore, it is not evaluated for signs of spam and will not be filtered. However, if mail is sent via the Andrew outgoing mail server but is not directly delivered to the Cyrus mail server it is scanned by the Andrew Spam Filter. This is the case if you use the Andrew outgoing mail server to send mail to a "user@cmu.edu" address. The mail is first sent to the cmu.edu server. The Cyrus mail server then receives the message from cmu.edu and the message will be evaluated and scored by the Andrew Spam Filter.
- Unless you selected the option to DISCARD spam, conditions set using your Accept list are checked before e-mail is filed into your spam folder. If an e-mail meets one of the Accept list criteria, it is automatically delivered to your INBOX regardless of its score.
- Likewise, conditions set using your Filter list are also checked before e-mail is delivered to your INBOX. If an address or domain has been added to your Filter list, the mail is automatically filed in your spam folder OR discarded, even if it has not been flagged as spam. (Unless the address is also on your Accept list. Unless you selected the option to DISCARD spam, the Accept list takes precedence over conditions set in the Filter list.)

You must perform maintenance on your INBOX.spam folder.

It is up to you to decide whether to delete or keep the mail filed in your spam folder. Spam is not normally highlighted as being "unwanted" mail. The Andrew Spam Filter has identified common aspects of spam and used this information to develop its rules. Because legitimate e-mail can sometimes be mistaken as spam, you should review the mail in your spam folder before deleting it. You must also regularly clean out your INBOX.spam folder to avoid filling your mail quota.

If you find that spam from a particular source is not being filtered, or that legitimate mail is inaccurately being filtered, customize your Accept list or Filter list to accommodate these domains or e-mail addresses.

Last Updated: 9/26/06